Learning-Based Image Rendering
Öğrenme Tabanlı Görüntü İşleme

Muhammed Kerim SOLMAZ[1,*], Mehmet SARIGÜL[2], Levent KARACAN[2]
[1]Department of Computer Engineering, Bilecik Seyh Edebali University, Bilecik, Turkey
[2]Department of Computer Engineering, Iskenderun Technical University, Hatay, Turkey
ORCIDs: 0000-0003-2071-293X, 0000-0001-7323-6864, 0000-0003-2764-5258
E-mails: mkerim.solmaz@bilecik.edu.tr, mehmet.sarigul@iste.edu.tr, levent.karacan@iste.edu.tr
*Corresponding author.

***Abstract*—Image rendering is essential study field for computer science, robotics, and augmented reality. In the last decade, the increase in the graphics processing power of computers and the widespread use of deep learning networks have led to deep learning networks being at the heart of the studies on image rendering. The use of deeper networks improves the visual representation ability of the trained models and gives them the ability to render high quality images. In this study, various information is given about subjects such as image rendering, obtaining 3D data from 2D data, 3D image rendering, differentiable rendering and recent studies on this subject.**

***Keywords*—computer vision; depth image; point cloud generation; differentiable rendering**

***Özetçe*—Görüntü işleme, bilgisayar bilimleri,robotik ve artırılmış gerçeklik gibi çalışma alanları için oldukça önemlidir. Son on yılda, işlem yeteneği ve işlem gücü artan GPU'lar sayesinde derin öğrenme ,yapay sinir ağları ve görüntü işleme konuları aynı çalışmaların içinde daha fazla kullanılmaya başlanmıştır.Artan işlem gücü ile birlikte daha derin ve kompleks modeller eğitilebilmektedir.Bu çalışmada görüntü işleme,2B veriden 3B veri elde etmek , 3B görüntü işleme ve türevlenebilir işleme gibi konular hakkında ve son zamanlarda bu konu ile ilgili yapılan çalışmalar hakkında çeşitli bilgiler verilmektedir.**

***Anahtar Kelimeler*—bilgisayar grafikleri; derinlik; nokta bulutu üretme; türevlenebilir işleme**

I. Introduction

In computer science, data is kept in matrices of different sizes. Image rendering aims to extract meaningful visuals from these matrices. Recent studies with image rendering, 3D rendering, and differentiable rendering in the computer science community are mentioned in this study.

The minimum number of coordinates required to specify a point on an object is called dimension. Lines are one-dimensional because they lie on a single axis. Visuals are 2-dimensional because a point on the image is specified with two coordinates, x, and y-axis. Three-dimensional objects are x,y and z it is specified with three coordinates. What distinguishes 2-dimensional and 3-dimensional objects from each other is that 3-dimensional objects are additionally specified with z coordinates. Three-dimensional space is the name given to the medium that has width, height, and depth information [1].

Datasets with known depth information and camera intrinsic parameters are more suitable for use for 3D studies. This is because they can be represented more accurately and easily in 3D space by utilizing depth information. By using depth information, the object can be represented in 3D space. Objects represented in 3D space can be transferred back to 2D space after the desired transformation processes are performed [2].

Rendering is the process of obtaining an image from 2D or 3D data. The resulting image is called a rendered image. It is one of the sub-topics of computer graphics and has an essential place in computer graphics. It is a method needed and used in many different sectors,not only in computer graphics but also in architecture, marketing, construction,3D mapping, video games, etc. Usually, a 2D render is obtained from 3D data. Information such as the geometry of the 3D data, camera perspective, and lighting must be known [3].

One of the missing aspects of the rendering process is that it is not possible to propagate back through the render output after the render output is taken. As a solution to this problem, the differentiable rendering [4] process has been developed. In the differentiable rendering process, it is possible to return over the render output.

II. Materials & Methods

A. The Image Depth and Depth Completion

In computer vision, the concept of depth is the information about the distance of an object. The depth map is the image that indicates the depths of all things in an environment. Today, it is possible to create a depth map with various methods (laser scanning devices, cameras that can measure depth, and similar tools). It is essential for studies. It is so important because it is much easier to represent 2D data with depth information 3D. Not all cameras capable of measuring depth give successful results. These cameras are sufficient to help get an idea about depth information, but they create some missing regions in the depth map for critical studies. These missing regions cause the 3D representation to be missing during the 3D representation for 3D studies and reduce the work's efficiency (Fig. 1). Due to this problem, various depth map completion algorithms have been developed [5]–[7].

Figure 1: Depth completion

These smart algorithms have been developed to complete the missing regions in the depth map. Depth maps completed using these algorithms are more suitable for 3D studies.

B. Point Cloud

A point cloud is a representation of an object as points in 3D space.Points are positioned according to depth information and camera intrinsic values in the 3D coordinate system. By using the all points in 3D space, the whole object can be observed. A point cloud is a very nice tool for examining and studying a 3D object. The object can be examined from all angles, and points can be manipulated according to the work to be done. It is possible to create a colored point cloud [8] in data with RGB-D (color and depth) knowledge. Nowadays, the process of creating a point cloud is relatively easy when enough data is available. Various algorithms are available for this process. These algorithms ensure that 3D data is represented with points (Fig. 2) [9], [10].

C. Differentiable Rendering

Image rendering is one of the cornerstones of computer vision. This is because data can be represented visually. The work area is pervasive and is used in almost all studies related to computer vision. Although it is used frequently, image rendering has some problems. At the beginning of these problems are the losses that occur when the 2D render output passes back to the 3D space. Since there is no back-propagation after rendering, it is impossible to return to the 3D space without loss. This problem has been solved with differentiable rendering [4]. With the differentiable rendering process, it becomes possible to return to the 3D space when necessary after rendering by relating the properties of the 3D scene. Differentiable rendering has become a very popular field of the study recently. It has been used in many studies.Some of these studies are, NeRF [11] work that makes 3D scene estimation by spreading back the data such as light, opacity from the 2D images, SynSin [12] is the estimation of the image that will be formed when the camera perspective of the 2D image is changed, Pytorch3D [8] work that creates colored point clouds in 3D space and returns from these point clouds back to 2D space, deep learning based 3D rendering work using Cuda cores [13]. It seems that its popularity will continue to increase.

III. Results & Discussion

These days, when the importance of learning-based methods is better understood, computer vision, deep learning, machine learning, and computer graphics have become fields of study where it is inevitable to use them together. Differentiable rendering strengthens the connection by increasing the relationship between these fields of study. Differentiable rendering will be the main subject of computer vision studies to be made in the future. The reason for this is that it allows learning-based studies because it pushes back-propagation possible. With this advantage, it is likely that studies on such as scene estimation, pose estimation, 3D visualization errors, and 3D optimization will increase. The development of these fields of work will allow the direct action of many industries (architecture, marketing, construction, 3D mapping, video games, etc.). Therefore, it will be a method that will raise many sectors, especially learning-based 3D rendering studies and computer vision studies, to much more advanced levels than they are now.

IV. Conclusion

As a result, differentiable rendering and learning-based image rendering should be studied for 3D studies in these times when computer vision is gaining importance. Anyone working in computer graphics, deep learning, or computer vision should know about differentiable rendering and learning-based rendering.

Appendices

Author Contributions

All authors equally contributed on writing the paper.

Figure 2: RGB-D image to point cloud

Acknowledgments

The study was supported by grant 120E447 from the TUBITAK.

Conflicts

None declared.

Ethical Declaration

This article does not contain any studies involving human participants and/or animals performed by any of the authors.

REFERENCES

[1] Kutlu Y, Alanoglu Z, Gokcen A, Yeniad M. Raspberry Pi based intelligent robot that recognizes and places puzzle objects. Journal of Intelligent Systems with Applications 2019; 2(1): 85-89.

[2] Yasar A, Uslu E, Cakmak F, Altuntas N, Amasyali MF, Yavuz S. Autonomous mobile robot navigation in structured rough terrain. Journal of Intelligent Systems with Applications 2018; 1(1): 67-74

[3] Uslu T, Cetin L, Gezgin E. Preliminary study of a surgical navigation with point based registration method. Journal of Intelligent Systems with Applications 2021; 4(1): 6-9.

[4] Kato H, Beker D, Morariu M, Ando T, Matsuoka T, Kehl W, Gaidon A. Differentiable rendering: A survey. arXiv preprint 2020; arXiv:2006.12057.

[5] Zhang J, Ramanagopal MS, VasudevanR, Johnson-Roberson M. Listereo: Generate dense depth maps from lidar and stereo imagery. In IEEE International Conference on Robotics and Automation (ICRA), May 31-August 31, 2020, Paris, France , pp. 7829-7836.

[6] Imran S, Long Y, Liu X, Morris D. Depth coefficients for depth completion. In IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), June 15-20, 2019, Long Beach, CA, USA, pp. 12438-12447.

[7] Zhang Y, Funkhouser T. Deep depth completion of a single RGB-D image. In IEEE/CVF Conference on Computer Vision and Pattern Recognition, June 18-23, 2018, Salt Lake City, UT, USA, pp. 175-185.

[8] Shinohara T, Xiu H, Matsuoka M. Point2color: 3D point cloud colorization using a conditional generative network and differentiable rendering for airborne LiDAR. In IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), June 19-25, 2021, Nashville, TN, USA, pp. 1062-1071.

[9] Zhou QY, Park J, Koltun V. Open3D: A modern library for 3D data processing. arXiv preprint 2018; arXiv:1801.09847.

[10] Luo S, Hu W. (2021). Diffusion probabilistic models for 3D point cloud generation. In IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), June 20-25, 2021, Nashville, TN, USA, pp. 2837-2845,

[11] Mildenhall B, Srinivasan PP, Tancik M, Barron JT, Ramamoorthi R, Ng R. Nerf: Representing scenes as neural radiance fields for view synthesis. In 16th European conference on computer vision, August 23-28, 2020, pp. 405-421.

[12] Wiles O, Gkioxari G, Szeliski R, Johnson J. Synsin: End-to-end view synthesis from a single image. In IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), June 13-19, 2020, Seattle, WA, USA, pp. 7467-7477.

[13] Ravi N, Reizenstein J, Novotny D, Gordon T, Lo WY, Johnson J, Gkioxari G. Accelerating 3D deep learning with pyTorch3D. arXiv preprint 2020; arXiv:2007.08501.